



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The VM is Written in Pahlavi Script

Citation for published version:

Herrmann, JM 2017 'The VM is Written in Pahlavi Script' pp. 1-12.

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



The VM is Written in Pahlavi Script

J. Michael Herrmann
School of Informatics
University of Edinburgh

1st August 2017

Abstract

The late medieval VM has resisted decryption and was considered a meaningless hoax or an unsolvable cipher. Here, we provide evidence that the VM is written in natural language by establishing a relation of the V alphabet and the Iranian Pahlavi script. Many of the V characters are upside-down versions of their Pahlavi counterparts, which may be an effect of different writing directions. Other V letters can be explained as ligatures or departures from Pahlavi with the intent to cope with known problems due to the stupendous ambiguity of Pahlavi text. While a translation of the VM text is not attempted here, we can confirm the V-Pahlavi relation at the character level by the transcription of many words from the VM illustrations and from parts of the main text. Many of the transcribed words can be identified as terms from Zoroastrian cosmology which is in line with the use of Pahlavi script in Zoroastrian communities from medieval times.

1 Why is V-ese difficult?

All writing systems in the world [6, 5] require some effort in acquisition and use. While for some groups of languages, difficulty and differences may be comparatively small [12], in others the complexity of the script can appear forbidding for all but a minority of scribes. Religious observance, for example, may require the adherents to continue using a script or language that no longer adapts to its language environment and that may thus tend to become ambiguous or incomprehensible. Diacritics or *matres lectionis* were devised as a remedy in order to retain a unique pronunciation and, supported by extensive commentaries, continuing understandability. However, such additional efforts may not be considered necessary, if the oral tradition in the community is sufficiently strong, such that the texts do not have to be extracted from the writing itself, but are rather remembered while being read. Here, we present evidence for the hypothesis that the VM is difficult to read because it derives from such a tradition.

One particular case of a language that is notoriously difficult to read is Pahlavi as it was used in medieval Zoroastrian scriptures, commentaries, and a few other texts [2]. Over the few

centuries of the language evolution, many Pahlavi letters have coalesced, for example, there is, in contrast to the older Avestan versions of the texts, only one letter for the phonemes *d*, *g*, *j*, and *y*. Moreover, letters are usually joined in Pahlavi script and can appear thus similar to other letters: E.g., in addition to its proper meaning, a letter can be indistinguishable from as much a sixteen different phoneme or letter combinations [8]. In some words, corrupted forms of letters have become a standard that is accepted to various degrees by the scribes. In addition to Persian words, Pahlavi contains also a large number of heterograms, i.e. around a thousand, partially very common words of Aramaic origin that are meant to be read in Persian (like the Latin abbreviation *i.e.* is read in English as *that is*). Finally, as for many other ancient texts, material decay, language drift, scribe errors, unfamiliarity with the original cultural context, and, possibly, the need of the writers to hide the content from contemporary hostility, also contribute to the difficulty of the translation of the text.

Concerning recent work on the VM, statistical approaches [1, 7, 10] that search for non-random features in data may be bound to fail if the target is quite random to begin with. The standard V character set (EVA) is not too helpful either, because it is unrelated to the phonemics, it breaks some of the letters into smaller parts, and fails to identify ligatures, all of which may further reduce the strength of the statistical analysis. In addition, the extensive 19th century literature dedicated to religious writing, see e.g. [11] was difficult to access until scanned copies became available online recently, and, finally, it may be construed that our academic habits thwart the systematic study of matters as obscure as the VM.

The VM which is written on over 200 vellum pages has been dated between 1404 and 1438 (University of Arizona, 2011), but its history is largely unknown until the discovery by the bookseller V in 1912. Apart from a few cautious attempts such as Ref. [3], so far no progress has been achieved in deciphering the VM nor even a decision was reached whether the VM has any meaningful content at all. In order to prove the hypothesis that the VM is a readable text with an interest in itself, we present here a comparison of the V and Pahlavi scripts (Section 2), provide evidence for the proposed relation between the two alphabets by a number of examples from VM illustrations as well as from its running text (Section 3). Finally, we will draw (in Section 4) some conclusions on the context in which the content of the VM may have originated.

2 Letters are reverted Pahlavi characters

Comparing the V-ese and Pahlavi scripts, we find that many of the characters are upside-down versions of each other, see Table 1. This may be due to the different writing direction of the two scripts. A similar effect that was observed also in the earlier sinistrodextral Brahmi script [4], in which also some of the letters appear as upside-down adoptions from its likely predecessor Aramaic (right to left). Pahlavi, that ultimately derives also from the Aramaic alphabet, has retained the dextrosinistral direction, while the VM is written in the opposite direction.

In this way, six of the about 20 V letters can be explained directly (*a*, *h*, *s*, *S*, *r*, in our notation, see Table 1, and *K*, see Table 2). Two more letters (*d*, *c*) differ from *s* and *S*,

respectively, only by an inverted breve diacritic. In addition, there are three more letters that obtain by rotation about a different angle (*t*, *y*) or by mirroring (*z*). The similarity of eleven out of the comparatively small number of letters of the two alphabets can be considered as a clear indication of a relation between V and Pahlavi (P). Below we will see that the relation extends also to the phonemic level. Two letters *o* and *n* that occur frequently in the VM, differ from their counterpart in the P alphabet. It is tempting to relate V *o* to P *pe*, but comparison to P's ancestor, Syriac, suggests rather an association to *waw*. This also supported by the frequent use of *o* as a word separator in the VM. In Pahlavi a vertical bar is used for this purpose, which is of similar shape as P *v* and may thus have lead to the transition to the more distinctive Syriac letter *o*. Further analysis of the V text will shown whether *o*, *y* or *a* also have a grammatical function. Based on the phonetic content (Sect. 3) of the letters, we assume that, in contrast to Pahlavi, the nasal alveolar is not part of the spectrum of V *o*. This position is given to V *n* in a supposed attempt at improved readability.

The remaining V letters are the “capitals” *B*, *K*, *M*, *P*, or occur only very rarely. Their shape may be due from a fusion of the respective Pahlavi characters with a vertical stroke (P word separator). We do not consider the capitals as ligatures, as they are used also within words or after the word separator (*o*). It is also interesting that the capitals and V *n* have a correspondence in shape and sound to (upside down) characters of the Modi script from the Maharashtra region.

Box 1: Comments on Table 1.

The letters are given in the order to the Aramaic alphabet with *resh* taking the place of phonetically similar *lamed*, and *jod* is placed near *daleth* and *gimel* with which it is interchangeable in Pahlavi. Frequently occurring corruptions are given in [brackets] [9]. Strokes from neighbouring characters are removed from the V letters.

[*a*] Appears in B-Pahlavi as a raised character. \aleph represents a glottal stop.

[*B*] We could not find enough evidence for systematic use of two variants (*B* and *H*) of this character.

[*g*] Occurs usually in final position, elsewhere *y* is used instead.

[*o*] V *v* resembles Syriac *vav* (ⲅ), psalter Pahlavi *vav* is identical to *resh*

[*y*] the letter represented here is *daleth*. The actual P-Pahlavi letter *yod* (Ⲋ) shows an interesting similarity to the inverted breve diacritic of V *d* and V *c*. Many words have an otiose *y* ending.

[*c*] This character occurs rarely in the manuscript such that the transcription is uncertain.

[*z*] is often (or easily) confused with *r*.

3 Vocabulary relates to Zoroastrian religion

V-ese and Pahlavi are not identical. By the introduction of a number of additional characters (*d*, *g*, *n* and initials), reading a V-ese text may be easier than the original Pahlavi. The history of the deciphering of the VM, does not support this claim. Analysing plant and star names, Bax [3] has suggested a similar reading for some but not all of the letters. We base our transcription on a larger number of samples from the manuscript and compare the

































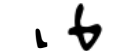
















V	B-Pahlavi	P-Pahlavi	transcription	name	#
			a, h, kh, ʾ	aleph	<i>a</i>
			b, a, h	beth	<i>B</i>
	 + 	 + 	d, t	daleth	<i>d</i>
			g	gimel	<i>g</i>
			y, g, d, [k, b]	yod	<i>y</i>
			v, o, u, [b]	waw	<i>o</i>
			h, r, n, w, l, [i, g]	he	<i>h</i>
	 + 	 + 	č	zayin	<i>c</i>
			k, g	kaph	<i>K</i>
			r, l	resh	<i>r</i>
			m, t	mem	<i>M</i>
			n	nun	<i>N</i>
			s, ī, [š, a, h]	samekh	<i>s</i>
			p, f	pe	<i>P</i>
			z, č, p, [g, d, y]	sahde	<i>z</i>
			š, t	shin	<i>S</i>
			t, s, [d, r]	tav	<i>t</i>

Table 1: V characters and initials together with variants of the corresponding Pahlavi letters (so-called Book- and Psalter-Pahlavi). The last column shows the notation used here. See Box 1 for comments.

V	Components	Phon.	#
	,	h, a, kh	<i>n</i>
	,	ā, h, a, kh	<i>m</i>
		sP	<i>sP</i>
		SK	<i>SK</i>
		sN, SN, sR?	<i>sB</i>
		sT	<i>sT</i>
		m, k + o	<i>ko</i>
		e	<i>cy</i>
		ā, h	<i>co</i>

Table 2: Main ligatures and letter combinations from the VM. Only part of the implied phonemes are given in the third column. The last column refers to the transliteration in Table 1. The two or three strokes of *n* or *m* have a similar functions as *h* in the final or penultimate position. Ligatures involving the letter V *S* (“table”) represent the succession of two consonants usually in the beginning of a word. In some cases it is *s* rather than *S* that is represented. While *sP* and *sT* are obvious from the vocabulary, the remaining combinations will have to be reconsidered. First part of *ko* occurs rarely if ever alone. This ligature can represent *m*, *q*, *h*, *r*, *mn*, *mv*, *mr*, *mN*, etc. The combinations *cy* and *co* appear to represent single phonemes in some cases, see Appendix B. All ligatures are copied from f37r, the components in the second column are from Table 1. Strokes belonging to neighbouring characters were removed.

results with names from the Zoroastrian cosmological scripture *Bundahesh* [13, 8], which was composed in the 11th century, and with general vocabulary [9], such that we arrive at a more complete and more reliable transcription that is based not only on the similarity of the letter shapes. The translations given below should not be expected to do justice to the VM text. They are solely included to provide evidence for the proposed transcription.

In the appendix, we show two sets of words from the manuscript. The first (App. A) gives the names of the zodiac symbols and the corresponding month names both of which were passed down in the *Bundahesh* [13, 8] in paragraphs II, 2 and XXV, 20, respectively. Based on the well known symbols shown in the centres of f70v1 – f73v, the identification with the Pahlavi names is straight-forward, except for the two pages f71r and f72r1 which show the same symbols (Aries and Taurus) as f70v1 and f71v, respectively. We cannot answer the question whether the two repeated signs do in fact represent the missing Capricorn and Aquarius. Because two words (on f72r2 and f72v2) are unreadable due to creases, we are left with 18 words that are identifiable to a reasonable degree of certainty.

In a similar way, it is possible to transcribe from the illustration on f69v most of the 28 lunar mansions that are also listed in *Bundahesh* II, 2. Because of the short and repeated Pahlavi names of the mansions, a unique correspondence was possible for only 20 of the mansions, such that we did not include it here. Interestingly, the 1247 stone representation of the Suzhou star chart (1193) that shows the 28 Chinese constellations has a “cartouche” title beginning with the ideogram for *sky* that can also be seen in a corrupted and reverted form on f1r of the VM. This is not implausible considering the continuous exchange between Persia and China in historic times.

The four words in the center of f67v2 are (with transcription) *zoahd* (*zohr*), *oBarao* (*bahr*), *zary* (*zōr*) and *natag* (*nihadag*). The translation yields the words *sacrifice*, *lot*, *power*, *foundation* [9] that appear, given a Zoroastrian parentage, semantically related. The words are grouped around a small square-shaped picture of a swirl-radiating star which could represent a sacrificial fire.

App. B includes a transcription of words from the beginning of the third paragraph of f1r. This sample is included not only to show that the Pahlavi transcription applies to the main text, too, but also to demonstrate the difficulty of a translation of the text, which has, however, been noted by all translators of Pahlavi documents.

In the illustration on f77v, we find the words *oBam yHat otBaNat orShNat oMot dhNy oMotor* which can be transcribed as *bīm duxt wad-baxt rēšinad mīh dēn wizār* and is translated word-by-word as *fear daughter unfortunately wounded: false religion explanation* [9]. We add this sample to point out that the VM does not necessarily propagate Zoroastrian beliefs, but may as well be a critical representation. This sample, nevertheless, suggests that the “nudes” pages (f75r – f84v) represent medieval medical content. While the representation of nude bodies is rare in such contexts, similar scenes appear in Mughal miniatures, where, however, an obviously erotic perspective is taken. The nudes in the VM apparently serve a merely illustrative purpose or show an ironic intent.

Further evidence for the proposed transcription can be obtained from the “colophon” (f116v). The last line of the short text contains the words *arar dccy* that are, in contrast to the seemingly Latin script on this page, clearly readable. We propose the transcription *xwar*

day, which would refer to the 11th day of the 10th month of the Zoroastrian calendar [9]. This information is possibly given also in the penultimate line, where we may identify several attempts to write the word *māh* (month) preceding what appears to be a Roman numeral. The line above has some similarity to the P phrase “book finished” and includes possibly the place of origin (Trebisonta?). We note this in order to increase the likelihood that f116v indeed shows a colophon, the more interesting question, however, whether the \cap character before the lacuna at the end of this line was originally the initial character of a year, cannot be answered without further analysis of the velum.

4 Discussion

We have not been presented more than a few words, which is mainly due to the inherent difficulties in reading Pahlavi. It should be easy to identify more potential Pahlavi expressions, but it may be possible as well that the text is written in a different idiom, such as Gujarati, whose Parsi dialect contains many words of Avestan origin, but is not an Iranian language.

It is striking that the manuscript does not contain any obvious religious symbolism (apart from the crucifix on f79v, which may well be a later insertion or indicate a critical perspective taken on the Zoroastrian content) nor any other culturally identifiable elements. However, the astronomical charts of the VM are related to the world of the Zoroastrian subculture in the middle East or South Asia. They do not show any awareness of (earlier?) Arabic astronomy, but seem to follow the cosmological view in the Bundahesh.

Finally, we want to emphasise that we have no evidence that the VM was produced in Persia (or perhaps even western India). More likely appear the regions around the Black Sea where an exchange between Persia and the Italian cities of Genoa and Venice took place around the presupposed time of the production of the VM. Our opinion that the content of the VM is meaningful does not exclude the possibility that it is still a hoax, namely in the sense that it was copied to be sold rather than read. In this process or by later action, foliae with critical content may have been removed to further obscure the origin of the manuscript.

Although the proposed transcription is still preliminary, it is now possible to find many of the VM words in a Pahlavi dictionary [9, 8] using Table 1. We do not aim at presenting a translation of the VM and can neither exclude a Mandaic or even Manichaic origin, as these religions share several features with Zoroastrianism. We are also unable to provide a more precise phonemic account at this stage, although some of the differences (e.g. between V *d* and P *t*) may allow for such discussions. It will require a substantial effort to provide a complete translation of the VM, unless the text turns out to have been passed down also from other sources. However, as the VM does not appear to be identical to any of the better known Zoroastrism-related scriptures or commentaries, its content may as well have an interest on its own.

Acknowledgement

The author acknowledges the use of the high-resolution scans made available by Jason Davies.

References

- [1] D. R. Amancio, E. G. Altmann, D. Rybski, O. N. Oliveira Jr, and L. d. F. Costa. Probing the statistical properties of unknown texts: Application to the VM. *PLoS ONE*, 8(7):00–00, 2013.
- [2] F. C. Andreas and K. Barr. Bruchstücke einer Pehlevi-Übersetzung der Psalmen. In *Sitzungsberichte*. Preußische Akademie der Wissenschaften, Berlin, 1933.
- [3] S. Bax. A proposed partial decoding of the V script, 2014.
- [4] G. Bühler. *On the Origin of the Indian Brahma Alphabet*. K. J. Trübner, 1898.
- [5] P. Daniels and W. Bright. *The world’s writing systems*. Oxford University Press, New York, Oxford, 1996.
- [6] C. Faulmann. *Das Buch der Schrift enthaltend die Schriftzeichen und Alphabete aller Zeiten und aller Völker des Erdkreises*. Verlag der Kaiserlich-königlichen Hof- und Staatsdruckerei, Wien, 2nd edition, 1880.
- [7] G. Jaskiewicz. Analysis of letter frequency distribution in the VM. In M. S. et al., editor, *Concurrency, Specification and Programming. Proceedings of the international workshop CS&P*, page 250, Pullusk, Poland, 2011.
- [8] F. Justi. *Der Bundeshesh*. F. C. W. Vogel, Leipzig, 1868. (Note that pages in the Pahlavi part are in incorrect order).
- [9] D. N. MacKenzie. *A concise Pahlavi dictionary*. Oxford University Press, London, 1986.
- [10] M. Montemurro and D. Zanette. Keywords and co-occurrence patterns in the VM: An information-theoretic analysis. *PLoS ONE*, 8(6):0066344, 2013.
- [11] F. M. Müller, editor. *The sacred books of the East*, London, 1879-1910. Macmillan and Co.
- [12] E. Paulesu, N. Brunswick, and F. Paganelli. Cross-cultural differences in unimpaired and dyslexic reading: Behavioral and functional anatomical observations in readers of regular and irregular orthographies. In N. Brunswick, S. McDougall, and P. de Mornay Davies, editors, *Reading and Dyslexia in Different Orthographies*, chapter 12. Psychology Press, 2010.
- [13] E. W. West. *The Bundahishn (“Creation”), or Knowledge from the Zand*, volume 5 of *Sacred Books of the East*. Oxford University Press, London, 1897.

A Zodiac pages (f70v1-f73v)

All descriptions were found within the V script around the margin (for f70v2, within the margin) of the central image that shows a depiction of the zodiacal sign.

★ Notes


✂ The centre image shows Aries in repetition of f70v1, but also the text does not show much evidence for the interpretation as Capricorn. The first letters of the V constellation are ignored, so the transcription is questionable.

⌘ The centre image shows Taurus in repetition of f71v, but also the text does not show much evidence for the interpretation as Aquarius. The first letters of the V month and V constellation are ignored, so the transcription is questionable. Alternative spelling: Ašwahišt

11200 1430

Alternative spelling: **Wahman**.

17621

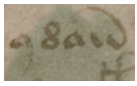
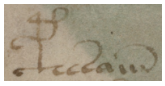


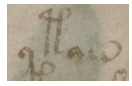
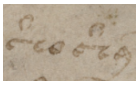


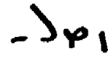
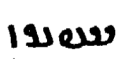

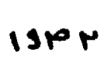
✂ Alternative spelling:  nemāsp.

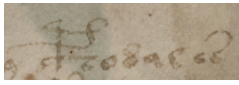
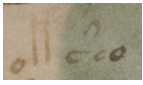
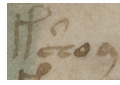
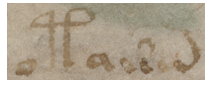
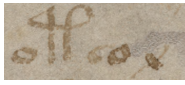
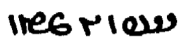



১১৬।


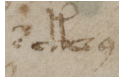
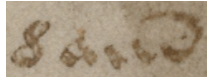


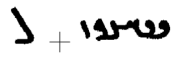

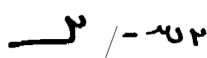
[illegible]

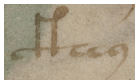
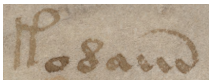


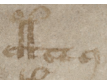



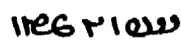
B First folio text (f1r)

Passage from the beginning of the third paragraph of f1r. Not all translations from [9] are shown. Our transliteration shows several inconsistencies, which may be due to the complexity and development of the Pahlavi language and will require further analysis. E.g. V *otr* retains P *t*, while V *dody* uses *d* for P *t* in accordance with the transliteration [9]. Final *o*, as in P for *dody*, is often ignored as an *otiose stroke* [9], see also the final character of *Spandarmad*. In VM, more often leading *o* are otiose, e.g. in *Spandarmad* after the line break, while in V *otr* the leading *o* is considered as part of the word.

V						
Vt	yNahr	{šP}(h)rahm	otr	{šP}hg	yMahn	dody
P						
Pt	gōhr	spram	wider-	spīg	yma/xōn	dūdag
E	nature	flower/fragrant herb	pass by	sprout	blood	family

V					
Vt	{šP}oN(a)thr-(o)Mdo		Mdoy/Adwg	oMahrhn	oMhot
P					
Pt	Spandarmad		ādūg	mārāg	nimūd[an]
E	Spenta Armaiti		capable	sensitive	guide

V				
Vt	oBahdoNan	r + {SB}hg	Nahm	doy
P				
Pt	wizihīdan/baxtan	man + sahīg	nām / wihān	dō / duš-
E	separation / distribute	me + worthy	name / cause	two/evil

V					
Vt	{ST}hy	BoNahm	{SP}g	{SP}on(a)tr-(o) Mdy	
P					
Pt	stī	bunih	spīg (?)	Spandarmad (?)	
E	being	beginning	sprout (?)	Spenta Armaiti (?)	

Notes

Curly brackets enclose {ligatures}. Square brackets indicate [inserted characters]. Round brackets indicate (ignored characters). A hyphen stands for a line break. Small strokes appearing in V either as *c* or *ι* are transcribed here as *h*, i.e. are considered to indicate a lengthening of the nearest vowel.

To explain the ignored V *a* in *Spandarmad*, the P *d* could be considered as a contraction of V *a* and V *t*.

The chapter on *The Nature Of Plants* (Bundahesh, Ch. XXVII) mentions *Spandarmad* [13]